THE ANGOFF METHOD AND RATER ANALYSIS: ENHANCING
CUTOFF SCORE RELIABILITY AND ACCURACY

THESIS

Presented to the Graduate Council of the
University of North Texas in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Charles E. Baker, B.S.
Denton, Texas
December, 1993

At times called a philosophy and other times called a process, cutting score methodology is an issue routinely encountered by Industrial/Organizational (I/O) psychologists. Published literature on cutting score methodology appears much more frequently in academic settings than it does in personnel settings where the potential for lawsuits typically occurs more often. With the passage of the 1991 Civil Rights Act, it is no longer legal to use within-group scoring. It has now become necessary for personnel psychologists to develop more acceptable selection methods that fall within established guidelines. Designating cutoff scores with the Angoff method appears to suit many requirements of personnel departments. Several procedures have evolved that suggest enhancing the accuracy and reliability of the Angoff method is possible. The current experiment investigated several such procedures, and found that rater accuracy methods significantly enhance cutoff score reliability and accuracy.
ACKNOWLEDGEMENTS

The number of people I would like to acknowledge is extensive, and I received help and encouragement from many quarters, particularly my friends and fellow classmates. However, I specifically would like to acknowledge my committee members and employees of the City of Fort Worth Personnel Division.

Drs. Doug Johnson and Mike Beyerlein have both been immeasurable sources of guidance and support throughout this project and during my Master's degree training in Industrial/Organizational Psychology. Dr. Bob Kelsey has been an excellent source of information in relation to experimental design and statistical analysis. To all three committee members, I extend a sincere, heartfelt thank you.

Dick Hodapp and his unit in the Testing Division of the City of Fort Worth Personnel Department are the other parties I wish to acknowledge. Dick was particularly helpful in pointing me toward a challenging research project that has both applied and academic implications. To him and others in his division, I extend my heartfelt thanks.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................ iii
LIST OF TABLES ........................................ v

Chapter

I. INTRODUCTION .......................................... 1

  Cutting Scores
  Selection Method History
  Legal Issues
  Methods of Cutoff Scores
  Enhancing Cutoff Scores
  Angoff Method
  The Raters
  Overview of Study

II. METHOD ............................................... 13

  Subjects
  Materials
  Procedures

III. RESULTS ............................................. 20

  Post Judgement Interview
  Internal Reliability of Raters and Judged Items
  Cutoff Scores
  Accuracy of Raters
  Rater Analysis

IV. DISCUSSION .......................................... 27

  Major Findings
  Implications of the Study
  Limitations of the Study
  Future Suggestions for Research

APPENDICES ............................................. 36

REFERENCES ............................................ 158
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Judges Cutoff Score Comparison</td>
<td>21</td>
</tr>
<tr>
<td>2. Item Judgement Accuracy: SME, TR, and UR Groups</td>
<td>24</td>
</tr>
<tr>
<td>3. Item Judgement Accuracy and Inter-rater Reliability for Total and Selected SME, TR, and UR Samples Based on Judge-Total Correlations</td>
<td>26</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

As long as the expectancy of success on the job is equal for two or more groups, the use of separate cutting scores is justified (Cascio, 1991). This statement is the essence of one side of the controversial issue of within-group scoring or as it is more commonly referred to "race norming." Hartigan and Widgen (1989) went even further when they recommended the reporting of a within-group expectancy score for each candidate with respect to the interpretation of scores on the General Aptitude Test Battery (GATB).

With the passage of the 1991 Civil Rights Act, Cascio's (1991) approach and similar ones have subsequently been rendered illegal. It is no longer legal for organizations to use within-group scoring to achieve goals outlined by their individual affirmative action plans. Thus, it has become necessary for psychologists in industrial settings to develop more acceptable selection methods that fall within the legal, ethical, and professional guidelines established when the new laws were enacted.

Cutting Scores

According to Cascio, Alexander, and Barrett (1988) as early as 1889 the Supreme Court in *Dent v. West Virginia*
upheld the denial of a medical license and noted that
difficult standards cannot be objected to unless the
procedure is not valid. They also discuss a more recent
case presented before the Supreme Court, Board of Regents of
the University of the State of New York v. Toronto (1980).
This case revolves around the denial of a chiropractic
license to an individual who had failed the exam seven times
by a very narrow margin, and the denial was upheld by the
court. Justice Stevens' remarks are typical: "Examinations
are a permissible method of determining qualifications, and
lines must be drawn somewhere," (p.493). The concept that
lines must be drawn is common in many areas of the law.
Numerous issues surround selection methods based on cutting
scores and to get a better perspective of how current
selection policies evolved, a historical overview is
warranted.

Selection Method History

Cascio, (1991) feels that psychologists' first law is
glaringly obvious: People are different. Physical and
psychological variability is all around us. As scientists,
our goal is to describe this variability, and through laws
and theories, to understand it, to explain it, and to
predict it. Measurement is one of the tools that enables us
to come a bit closer to these objectives. In a personnel
context, these measurements have significant economic and
legal impact upon both employee and employer.
Various cutoff models exist, but the two primary ones are single stage (non-sequential) and multiple hurdle (sequential). Non-sequential cutoff selection methods utilize the multiple regression or multiple cutoff method. According to McKinney and Collins (1991), in municipalities as in other governmental organizations, the multiple hurdles and its variations are the most commonly utilized selection methods. The remainder of this investigation will limit itself to the topic of the sequential selection method due to the investigator operating in a public organizational environment.

No organization wants to lose productivity due to the selection and placement of inferior workers, and some companies attempt to follow a "correct social" agenda by employing members of protected groups. Seeking to reach the middle ground, companies look for the best of both worlds. Although no one selection procedure provides everything for organizations, the cutoff method appears to be an acceptable approach.

McKinney and Collins (1991) investigated three popular personnel selection methods. Their study compared the top down raw score approach, the minimum cutoff procedure, and the within-group standardized score method. Their study findings suggest the cutoff scoring procedure produced the second highest gain in productivity and served to allow the most equitable mechanism for extending preferential
treatment to protected groups while also reducing the potential for reverse discrimination.

Although cutoff score methodology appears to be an acceptable selection method compromise, it has both positive and negative facets associated with the process. Cascio (1991) states that raising the cutoff score \(X_p\) decreases the probability of erroneous acceptances, but it simultaneously increases the probability of erroneous rejection. Lowering the cutoff scores has exactly the opposite effect. Another instance noted by Cascio (1988) where cutoff scores were raised extremely high deals with the case of Rogers v. International Paper Co. (1975). The cutoff scores were set so high on certain tests that 40% of the incumbent skilled craftsman would not have been able to gain admission to the craft.

Martin and Raju (1992) discuss two other effects of raising cutoff scores. First, as a result of being more selective, the expected performance level of selected applicants increases. Second, being more selective means that more applicants must be processed to fill a fixed number of openings, usually the case for government positions, which increases hiring costs.

Legal Issues

Title VII of the Civil Rights Act of 1964 laid the groundwork for many court cases to attempt interpretation of the social equity issues that evolved as a
result of its passage. Since its inception, conflict has existed between those who consider the intent of the Civil Rights Act as insuring equal results for protected groups now, and those who argue that the legislature means equal opportunity for all individuals forever (McKinney & Collins, 1991).

Cascio et al. (1988) cite several decisions by the Supreme Court that have supported and helped to define the role cutoff scores play in personnel selection. Griggs v. Duke Power Co. (1971) was the case where the Supreme Court first addressed cutoff scores within the context of Title VII. In the body of the opinion, the court makes very little mention of cutoff scores, except to note that the passing scores on the Wonderlic and the Bennett were very stringent criteria, since they were set approximately at the median score achieved by high school graduates.

Another case cited by Cascio et al. (1988) as a landmark decision was Albermerle Paper Co. v. Moody (1975). A poorly devised and implemented validation procedure involved testing some individuals at the top of a job progression. The court pointed out that favorable scores above some particular cutoff score for those individuals did not mean that the test was a good measure of minimum qualification for workers in lower level jobs.

Later, a more conservative Supreme Court influenced organizations with a decision that required plaintiffs in
discrimination cases to provide more evidence to prove their cases. McKinney and Collins (1991) point to the *Wards Cove Packing v. Atonio* (1989) as a Supreme Court ruling that some civil rights leaders felt has inaugurated a period of retrenchment and reversals in the effort to rectify past discrimination. The court ruled that minorities should not be favored in selection results at Alaskan salmon canneries by saying that the plaintiff bore the burden of proof when disproving an employer's assertion that the adverse employment action was based solely on legitimate, neutral considerations.

With the passage of the Civil Rights Act of 1991 several changes occurred that will have a significant impact on personnel management practices. The primary change was the current Civil Rights Act reversal of seven Supreme Court decisions, all of which were largely favorable to employers, thus making it easier for applicants and employees to challenge as discriminatory any type of employment decision, including hiring, promotion, layoffs, and termination (Carey, Mazaroff, & Tranchida, 1992).

Uinick (1992) points to an interesting adjustment made by the Civil Rights Act of 1991. It is no longer legal to adjust the scores of, use different cutoff scores for, or otherwise alter the results of, employment related tests on the basis of race, color, religion, sex, or national origin. The law does however, provide for banding of scores, which
by itself is completely blind to race and gender. It does no more than to determine that actual scores are statistically equivalent, thus eliminating the potential for statistical error of measurement.

**Methods of Cutoff Scores**

There are two major approaches to establishing cutoff scores. First, the systematic norm-referenced strategy that appears in the selection context is essentially a human resource planning approach. Information regarding projected personnel needs, the history of the proportion of offers accepted, and a large-sample distribution of applicants' test scores are integrated to set a cutoff point on a test that will yield the number of new hires needed (Cascio et al., 1988).

Second, Cascio (1988) describes content-related validity settings as the method most of the research literature on cutoff scores target. The literature refers to cutoff scores as "criterion referenced cutoff scores"; however the "criterion" referred to is a subjective standard that exists for the person who is providing judgmental inputs to the system. One of the more popular criterion-related procedures is the Angoff method.

**Enhancing Cutoff Scores**

Berk (1986) thoroughly reviewed the research literature on several methods for setting performance standards. Of the several methods studied, the technique advocated by
Angoff (1971) was the most popular. He also concluded that it had the best balance between technical adequacy and practicality.

**Angoff Method**

Cascio et al. (1988) felt that of three similar criterion-related cutoff methods, the Angoff procedure is generally preferred on the grounds that it is at least as reliable as either of the other two, and it requires substantially less rater time and effort. This becomes especially important in a budget controlled environment.

Maurer, Alexander, Callahan, Bailey, and Dambrot (1991) describe the Angoff method as a technique where judges rate the probability that a barely or minimally competent person would answer each test item correctly. The standard or critical score becomes simply the average of the item proportions multiplied by the number of items (or simply the sum of the item proportions). The judges are typically subject matter experts (SME’s) and thus, this method has intuitive appeal in that experts from the relevant field use their expertise and knowledge to help determine minimum performance standards. As a result, it may be better received by various constituencies than a standard that is set arbitrarily or based solely on consultants’ judgement.

Cascio et al. (1988) points out four disadvantages seen when utilizing the Angoff or modified Angoff procedures. One, all have very poor inter-judge and alternate methods
reliabilities. Two, it is quite time consuming and difficult to administer. Three, there is no consensus (nor even much discussion) as to whether items with extremely low inter-judge reliabilities should be eliminated from the test, eliminated from consideration when setting the cutoff score, or if such differences simply should be ignored. Finally, all of the methods seem to suffer from a lack of acceptability from outside stakeholders. While persons without a background in psychometrics may be willing to accept that experts could make such judgements regarding individual test items, they have difficulty comprehending how such judgements can lead logically to a single cutoff point.

The Raters

There have been many investigations in educational settings on the accuracy and integrity of cutting scores using the Angoff method. One study, Norcini, Shea, and Kanya (1988), recorded findings that suggest the standard-setting process can be improved by providing performance data for items to prevent unrealistic standards.

Cascio et al. (1988) suggested that the Angoff method can be improved by doing a rater analysis. This analysis would be similar to the item analysis performed on test items (that is, rater-total correlations analogous to item-total correlations, and so forth), and to eliminate raters on much the same basis that items are eliminated.
The rater analysis suggestion by Cascio was the area the investigator's experiment explored. According to Cascio et al. (1988) the technical and practical problems associated with setting cutoff scores are common problems for personnel psychologists. He also states that this issue has received a great deal of attention in the educational measurement literature, but very little discussion of it has appeared in the personnel or I/O psychology literature, despite its importance.

Maurer et al. (1991) performed an experiment in an educational setting utilizing Cascio's suggestion to perform a rater analysis. In the current experiment, the investigator replicated the Maurer et al. study, except the experiment was performed in an industrial setting. The applied and logistical issues involved in performing the proposed experiment in an industrial setting necessitated some procedural and statistical alterations as noted.

Overview of Proposed Study

The Angoff method has been utilized extensively in educational arenas; however it has had limited application in personnel type settings. Most organizations look to methods that are cost effective, legal, and require a minimal number of man hours to get a more than adequate return on their investment.

Taking these factors into consideration, the current experiment investigated several Angoff accuracy techniques
advocated by Maurer et al. (1991) with minor alterations to accommodate the dynamic and specialized environment associated with most organizations. In this particular organization, the selection method utilized is the multiple hurdles process and proceeds in the following manner. The applicant completes an application for a clerical position, and if they meet minimum qualifications (MMQ's) they are allowed to move to the next hurdle, that is, the clerical test. After passing the clerical test, they are placed in a competitive pool of applicants who are eligible for further employment consideration.

Maurer et al. (1991) stated two purposes for their study: (a) to explore the larger issue of the effect of expertise on judgement using the Angoff method, and (b) to illustrate and examine psychometric principles and methods that can be used in criterion referenced testing, where accurate and reliable cutoff scores are important. Additionally, the Maurer experiment utilized a classroom test that was standardized in content and administration. The present investigation had similar purposes but was conducted utilizing an employment selection test, The General Clerical Abilities Exam, developed by the City of Fort Worth Testing Division.

In the experiment done by Maurer et al. (1991), generalizability theory was utilized to make a detailed comparison of the reliability of the expert and non-expert
group. In the present study, the personnel setting dictates that a more time efficient and cost-effective method should be employed. The Cronbach alpha formula gave an accurate comparison of the subject matter experts (SME) trained rater (TR), and untrained (UR) control groups. The current experiment is similar to the Maurer et al. study except it is targeted specifically to personnel selection issues where cutoff scores are especially important for productivity and prediction purposes. The expectation of the current study is that the SME group will have significantly more accurate rating indices than either the trained rater or an untrained control group. Additionally, the trained raters should have significantly more accurate scores than the control group.
CHAPTER II

METHOD

Overview

The current study was conducted utilizing archival test data obtained from a clerical test database maintained by a major municipality's testing division. Testing division personnel supervised the exam that was standardized in both content and administration. The division administrator of the testing department developed the exam and supervised testing personnel. Selected professionals from the testing division made recommendations regarding the subject matter experts (SMEs) who evaluated the test in the study.

Subjects

There were three groups of judges in this study. The first experimental group consisted of five supervisors considered subject matter experts. These SMEs were females age 31 to 45, and they are considered experts because of their previous employment at the position being evaluated and their minimum of three years of clerical supervisory experience. The judges were from five different departments within the City of Fort Worth, Texas.

The second experimental group, the trained raters (TR) and the control group, hereafter referred to as untrained
raters (UR) consisted of 20 judges each, and were selected randomly from students enrolled in undergraduate psychology courses. These students ranged in age from 18 to 32 years and contained 19 males and 21 females. To reduce generalization of clerical work experience during the Angoff judging task, any subject in the TR or UR groups who indicated clerical supervisory and/or more the two (2) years of clerical experience were rejected as raters. This procedure should present a less contaminated "job bias" group of TR and UR judges.

Materials

One exam, evaluated using the Angoff method, consisting of 125 items and having one overall score, was given to each validation study applicant competing for a clerical position at numerous departments throughout the municipality. The exam was developed by the Testing Division Administrator at the City of Fort Worth, Texas. After utilizing the Comprehensive Occupational Data Analysis Program (CODAP) to perform a job analysis, job requirements were determined and the nine subtests were used to evaluate relevant job abilities. The General Clerical Abilities Exam was a composite taken from over 30 other municipality exams. A factor analysis was done to identity the basic aptitudes necessary for the job and eliminating or reducing some test items and subtests. The analysis determined nine factors that subsequently became the exam subtests. The subtests
were selected based on factorial validity and practical validity. Item difficulty and adverse impact analyses were performed to remove any questionable items.

This exam is considered a power test measuring general clerical abilities. The exam consists of nine subtests: (a) Spelling (b) Vocabulary (c) Office Practices (d) Filing (e) Forms Checking (f) Reading Comprehension (g) Checking (h) Language Usage, and (i) Mathematical Reasoning. The tests were machine scored by computers. Machine scoring and utilizing archival/shelf data stored in personnel data banks allowed the individual test items’ difficulty to remain unknown.

In 1989, utilizing information maintained on the 268 subjects in the study, the exam was validated using archival data in addition to supervisor rating one year post-exam. Test validity (.31) was determined by concurrent validation and the range of actual item difficulty levels of .09 to 1.0 were established. Content validity was also built into the test and was based on the job and factor analyses.

Procedure

The following procedure is modeled after the Angoff method where raters are to estimate performance of person identified as being "minimally competent." These estimates are to be based on some job relevant criterion that in the present study will be entry level job performance for a clerical position.
Archival data drawn from the validation subgroup of minimally competent Clerical Worker I (criterion group) was the information utilized to compare the three groups. Information obtained from this municipality was provided by data maintained for employee selection of the Clerical Worker I position. The 23 minimally competent employees were identified by supervisor job performance rating derived from a pool of 25 individuals. Two individuals were removed for invalid performance review ratings.

A percentage sheet with values ranging from 10% to 90% in ten unit increments was provided to the SME's for completion. Written and oral instructions were given to the raters. The written instructions consisted of (a) the definition of a minimally competent person, (b) encouragement of raters to develop mentally a "minimally competent clerical employee prototype," and (c) encouragement of raters to determine mentally, what makes for marginal employees based on critical incidents, as they were evaluating each item.

The judges were asked to estimate, for each individual item on the 125 item test, the probability that a person who performed at a minimally acceptable level, in this entry level clerical position, would choose the correct answer. "Minimally acceptable employee" was defined by performance on the criterion measure (evaluation based on supervisors' standardized performance appraisals).
Individually, the SMEs were required to complete the Angoff ratings based on experience relative to their own clerical work units. They were then required to complete the Clerical prototype form individually. This form required performance information of both a specific and general nature. The specific performance standards were (a) quantity of task, (b) quality of task, and (c) and the time required to complete the requested task. The general performance standard required information describing KSAs and also tasks a minimally competent Clerk I was not capable of successfully completing.

As a group, the information the SMEs generated from the individual completion of the Clerk I prototype was discussed and a composite was developed. Following the completed exercises, the SMEs were interviewed about their responses to the questionnaire and the Angoff ratings.

The composite prototype developed by the SME's was presented to the TRs as a training tool. The TR group received the same directions as the SMEs for completing the Angoff rating. Further instructions to this group included (1) use this prototype description and the information on the percentage sheets as a frame of reference to make Angoff ratings (2) take one half (1/2) hour to study the prototype form, and (3) use the Clerk I description throughout the ratings to match subtest questions to described job functions.
The control group received no information about a prototype. The only information provided to this group was the description of a Minimally Competent Clerk I and the KSAs listed on the percentage sheets. Instructions were provided on how to complete the percentage sheets, but no instructions were given on how to determine a Minimally Competent Clerical Worker I. Judges in each group were interviewed after completion of the exercise. This procedure was meant to gather process information related to how the judges arrived at their Angoff ratings.

The criterion sample data was derived from a subset of the 268 subjects in the original validation sample. Archival test data were retrieved and compared to the criterion measure to identify employees in the validation study who performed at the minimally acceptable level. Data from their performance on each of the 125 items were also obtained. Thus, the normative data were available for each item, based on the performance of 23 employees who were identified as minimally competent. Item difficulty values, calculated on these employees for individual questions, served as estimates of "true" probability values for the individual items.

Two accuracy indices were determined. The first measure was an absolute deviation index that reflected how close each judge's assigned probability values and the true probability were. The Item Distance Accuracy (IDAcc)
method, summed across all items, the absolute deviation of a judge's assigned probability from the true probability for each item to compute a judge’s score.

The second index was the correlational accuracy index (rAcc) and reflected the correlation between each judge's assigned probability values and the true probability values. A rater total correlation was also computed. This correlation is similar to an item analysis and indicates the relationship between the judge's ratings of the items and the ratings by the rest of the sample of judges. These three procedures are psychometric methods advocated by Maurer et al. (1991).

Cutting scores were compared for the control, TR, and SME groups to determine if any differences that may exist could affect the rate of True Negatives, True Positives, False Negatives, and False Positives observed in the subjects studied.
CHAPTER III

RESULTS

Post Judgement Interview

Random polling of raters in the control group revealed that some judges based their ratings on their own perceived abilities, some on their perception of what they felt a minimally competent Clerical I employee could accomplish, and others simply followed the KSA guidelines provided with the Angoff rating form.

Interviewing the trained raters and the SMEs, revealed that both groups utilized the specific information provided to them by the investigator. Separately, the SMEs followed personal experience in making their ratings, and the trained rater group drew information from the prototype profile they were provided.

The majority of the 45 raters felt that the test had a great deal of face validity when compared to the provided information, and most judges completed the task between 40 and 55 minutes. The TR group was allowed 30 minutes to peruse the Prototype profile and when asked, none had questions prior to beginning the judging process.

Internal Consistency of the Test

The clerical exam validation sample used for the Angoff ratings exhibited an alpha of (KR 20) .91 extracted from the
268 subjects and the 125 item exam. The criterion sample from which the True Scores were obtained was based on the 23 clerical workers identified as minimally competent by their supervisor's performance ratings. The alpha obtained from this group was .89 based on the 125 test items.

Cutoff Scores

The cutoff scores presented by the three groups were 81.47 for the SME's, 81.73 for the TR's, and 81.91 for the UR group. No significant difference between groups was evident for cutoff scores $F (2,42) = .005, p > .05$.

Table 1

Judges Cutoff Score Comparison

<table>
<thead>
<tr>
<th></th>
<th>CUTOFF</th>
<th>STD.DEV</th>
<th>MNCUT</th>
<th>MXCUT</th>
<th>%CUT</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSd</td>
<td>81.19</td>
<td>13.93</td>
<td>56</td>
<td>112</td>
<td>.64</td>
<td>23</td>
</tr>
<tr>
<td>SMEe</td>
<td>81.47</td>
<td>9.27</td>
<td>71.60</td>
<td>94.73</td>
<td>.65</td>
<td>5</td>
</tr>
<tr>
<td>TRf</td>
<td>81.73</td>
<td>10.97</td>
<td>65.70</td>
<td>98.40</td>
<td>.65</td>
<td>20</td>
</tr>
<tr>
<td>URg</td>
<td>81.91</td>
<td>7.41</td>
<td>66.73</td>
<td>97.28</td>
<td>.66</td>
<td>20</td>
</tr>
</tbody>
</table>

aMN CUT = minimum cutoff score assigned by each group.
bMX CUT = maximum cutoff score assigned by each group.
c%CUT = percentage of correct items.
dTS = true scores computed based on criterion sample of minimally competent Clerk I workers.
eSME = Subject Matter Experts; TR = Trained Raters; UR = Untrained Raters.
Further examination of Table 1 shows more variance in the TR group than either the SME or UR group. The SMEs set a higher minimum cutoff score (71.60) and a lower maximum cutoff score (94.73). Comparison of cutoff scores by group had no effect on the number of True Positives, True Negatives, False Positives, nor False Positives.

**Accuracy of Raters**

Maurer et al. (1991) advocated two indices of judges rating accuracy. The first, the Item Distance Accuracy Score (IDAcc), was computed by summing across all 125 items, the absolute deviation of a judge's assigned probability from the true probability for each item. This index reflects how closely the judges assign probability values that match the true probability. The second index advocated by Maurer et al. (1991), the Correlational Accuracy Index (rAcc) was determined for each judge. A correlation between each judge's assigned probability values and the true probability values, this index indicated a judge's sensitivity to the true item difficulty levels. Subsequent analysis with this index involves transforming into Fisher's Zs.

Table 2 indicates that significant effects were found for the IDAcc $F(2,42) = 6.470, p < .005$ and rAcc $F(2,42) = 7.528, p < .005$. Post hoc (Tukey) examination of these two indices show significant differences occurred between the SME and TR groups, and the TR and UR group for both
measures. These differences were significant at the \( p < .01 \) level. The numbers next to IDAcc in parentheses beside the means are the average absolute deviations from the true item ps (in other words, the sum of IDAcc divided by 125. Higher IDAcc scores indicate less accurate ratings due to greater deviations from true scores. The opposite is true for rAcc, as the larger this index the more accurate the ratings due to greater sensitivity to true item ordering. The correlations between IDAcc and rAcc were negative for all three groups with the SMEs exhibiting \( r = -.50 \) (\( p > .05 \)), TRs \( r = -.71 \) (\( p < .001 \)), and the UR raters \( r = -.66 \) (\( p < .01 \)). According to Maurer et al. (1991), a negative correlation indicates that the more sensitive the raters were to criterion item ordering, the more closely they came to judging the true item ps and vice versa.

Rater Analysis

Another procedure performed by Maurer et al. (1991), the rater-total correlation was also computed. This procedure involves computing a correlation between a judge’s rating of each of the 125 items with the corresponding group (sum) rating of the item. An individual judge’s rating of each item was not included in the sum, in order to eliminate shared variance due simply to the judge’s own ratings. These correlations were transformed into Fisher’s Z’s, a mean Z was derived, and the mean Z’s were transformed back to correlations. Significant differences
Table 2
Mean and Standard Deviations of Judges Item Accuracy

<table>
<thead>
<tr>
<th></th>
<th>SME (n=5)</th>
<th>TR (n=20)</th>
<th>UR (n=20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDAcc Mean</td>
<td>23.53(.19)</td>
<td>27.58(.22)</td>
<td>24.94(.20)</td>
</tr>
<tr>
<td>SD</td>
<td>1.08</td>
<td>3.49</td>
<td>2.26</td>
</tr>
<tr>
<td>rAcc Mean</td>
<td>.36</td>
<td>.18</td>
<td>.32</td>
</tr>
<tr>
<td>SD</td>
<td>.05</td>
<td>.15</td>
<td>.13</td>
</tr>
</tbody>
</table>

IDAcc = sum across items of the absolute deviations of assigned ratings from true probability values; rAcc = correlation between true item probabilities and judge's ratings.

Numbers in parentheses are mean absolute deviations from true scores of assigned item probabilities.

were not found between for the rater-total correlations (F (2,42) = 1.540, p > .05), where the SME mean correlation was (r = .14), TR (r = .42, SD = .27) and UR (r = .53, SD = .20).

Based on the judges rater-total correlations, the judges with nonsignificant rater-total correlations were dropped from each group to compare the effect on cutoff scores and accuracy. This resulted in the SME group having n = 4, TR group n = 17, and the control group n = 19. This procedure also produced no significant difference in mean rater-total correlations.

A correlation was calculated between the rater-total correlations Fisher's Z for each judge and the rater's
corresponding item judgement accuracy values (IDAcc and rAcc) in all three groups. The SME's correlation was $r = .31 (p > .05)$ for the IDAcc, and $r = -.76 (p > .05)$ for rAcc. The TR's correlation was $r = -.33 (p > .05)$ for IDAcc, and $r = .11 (p > .05)$ for rAcc. The UR group correlation was $r = .19 (p > .05)$ for IDAcc, and $r = .07 (p > .05)$ for rAcc.

These data suggest that eliminating judges based on rater-total correlations would not be a useful technique to increase reliability or accuracy. However, to explore the effects of eliminating raters with nonsignificant rater-total correlations, the investigator removed three raters from the TR group and one rater from the UR group. Additionally, one SME rater with a low rater-total correlation, although statistically significant, was removed for comparative purposes. Slight and inconsistent differences were observed across measures for IDAcc, rAcc, and Cutting Scores for each group. As represented in Table 3, eliminating raters based on their rater-total correlation had minimal effects on the judge's IDAcc and rAcc, but in no case did the indices show a gain of more than three percent.
Table 3

**Item Judgment Accuracy and Interrater Reliability for Total and Selected SME, TR, and UR Groups Based on Rater-Total Correlations**

<table>
<thead>
<tr>
<th>Group (n)</th>
<th>M IDAcc&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SD IDAcc</th>
<th>M rAcc&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SD rAcc</th>
<th>CS&lt;sup&gt;c&lt;/sup&gt;</th>
<th>CS SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SME (5)</td>
<td>23.53 (.18)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.08</td>
<td>.36</td>
<td>.01</td>
<td>81.47</td>
<td>9.27</td>
</tr>
<tr>
<td>SME (4)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>23.63 (.18)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.22</td>
<td>.34</td>
<td>.01</td>
<td>83.93</td>
<td>8.59</td>
</tr>
<tr>
<td>Accuracy change</td>
<td>- .4%&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td>-.6%&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TR (20)</td>
<td>27.58 (.22)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.49</td>
<td>.18</td>
<td>.15</td>
<td>81.73</td>
<td>10.97</td>
</tr>
<tr>
<td>TR (17)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>27.44 (.21)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.73</td>
<td>.18</td>
<td>.16</td>
<td>79.12</td>
<td>9.67</td>
</tr>
<tr>
<td>Accuracy change</td>
<td>.05%&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td>0%&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UR (20)</td>
<td>24.94 (.20)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.26</td>
<td>.32</td>
<td>.13</td>
<td>81.91</td>
<td>7.41</td>
</tr>
<tr>
<td>UR (19)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>24.84 (.19)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.28</td>
<td>.33</td>
<td>.13</td>
<td>82</td>
<td>7.66</td>
</tr>
<tr>
<td>Accuracy change</td>
<td>.4%&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td>.3%&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>IDAcc = sum across items of the absolute deviations of assigned ratings from true probability values; rAcc = correlation between true item probabilities and judge's ratings.

<sup>b</sup>Numbers in parentheses are mean absolute deviations from true scores of assigned item probabilities.

<sup>c</sup>CS = cutting score.

<sup>d</sup>1 judge with significant rater-total correlation eliminated.

<sup>e</sup>3 judges with nonsignificant (P < .05) rater-total correlations eliminated.

<sup>f</sup>1 judge with nonsignificant rater-total correlation eliminated.

<sup>g</sup>Index's improvement/loss due to rater elimination process.
CHAPTER IV

DISCUSSION

The present study investigated the Angoff method for establishing cutoff scores and several psychometric procedures that might increase rater accuracy and reliability. The issue of person hours balanced against quality of work usually surfaces in any organization concerned with increasing productivity. The premise of this investigation was to evaluate whether raters could be trained to increase their cutoff score accuracy. Utilizing trained raters would be less expensive than pulling SME's from their jobs and should allow for more accuracy with cutoff scores than those set arbitrarily. This expectation was not met.

Findings

IDAcc and rAcc methodologies were supported in the current study suggesting rater accuracy techniques are useful in personnel settings. However, the usefulness of the rater analysis or rater-total correlation technique was not supported.

Significant differences were found between the SME and the TR groups, but no differences existed between the SME and UR group. Though puzzling, these data may have several explanations. Performance by the SME group, (although not
significantly different from the URs), with the IDAcc and rAcc indices showed greater accuracy by the SME group. Given the small n of the SME group, it would appear that statistical power was not adequate to result in a significant difference between the groups.

Performance by the UR group was surprising, especially considering the limited information they received about essential job functions of clerical workers. Their IDAcc and rAcc indices did not show quite the accuracy of the SME group, but there was no significant difference. There were however, significant differences between the UR and TR groups for the IDAcc and rAcc indices.

The TR’s performance was poorest of all the groups. Significant differences existed between the TR group and both the SME and UR groups. Further, these differences occurred with both the IDAcc and rAcc indices. This finding, although the opposite of the experiment’s expectations, suggests several dynamic processes may have contributed to these results.

First, the job being rated was an entry level, non-technical position. Many people have had contact with clerical workers and may have relatively accurate ideas about the job content of that position. Providing the raters with KSAs and essential job functions may have been all that was necessary to accurately judge items. Job content of high-technical and non-entry level positions are
much more complex and would probably lead to different results.

Second, logistics related to conducting applied research, such as maintaining test security, scheduling of SME and testing department personnel, led to the collection of the data erratically. Collection of the data was performed across a period of several months. Security measures for this exam did not allow for participation of numerous subjects, and therefore only ten subjects or less per session completed the Angoff ratings. Other security measures beyond the ten subject limitation, resulted in the temporal discrepancies of the data collection.

The UR group came from a sample of students who participated in the experiment early in the semester. The TR group came from a sample of students who participated in the experiment later in the semester, closer to finals. Motivation of the subjects in the study becomes a question, as the earlier subjects may have been participating due to scientific interests and extra credit points, whereas later subjects may have been motivated primarily for extra credit points to enhance their grades, with less regard for scientific interests.

Third, too much information may have been given to the TR group. The prototype presented to the TR group was very detailed and contain some jargon specific to the clerical position.
Questions about the judgement task arose in this study concerning the strategy the raters utilized during the Angoff ratings. The similarity of cutoff scores between the groups would suggest that the item responses hovered around a value of 80%. However, as Maurer et al. (1991) found in their study, the SME data in the present study were more accurate, but their accuracy indices and cutoff scores showed inconsistent variance. The inconsistent variance can perhaps be explained by the small \( n \) of the SME group and the different tasks required of clerical workers in the five different departments.

In their study, Maurer et al. (1991) suggest that if raters were hovering around a value such as 70%, when assigning probability values, it was done because they would have been conscious of the criterion (e.g., Cs are approximately 70%, etc.). Several factors in the present study suggest this phenomenon did not occur. First, the Maurer et al. study utilized SMEs who were graduate students and could have been aware of the 70% cutoff criterion for a minimally competent student. The current study utilized SMEs in the public sector, and results of the clerical exam were not released to anyone outside the personnel department; Therefore, they had no reference point on which to base their ratings. Additionally, the probability values for the true scores based on the criterion group's performance ranged from .09 to 1.00.
Second, the SMEs had more consistent item ordering than the TR or UR groups; however the reliability and item ordering of the UR group was of higher quality than the TR's. The prototype presented to the TRs consisted of two parts. The first part described specific performance measures of quantity of task, quality of task, and time required to complete the task. The second part described general abilities that would prevent a clerical worker from being successful as a minimally competent Clerk I, and was more subjective in nature. The subjective evaluation material in the prototype could help explain the poorer performance of the TRs when compared to the UR group.

Third, the SME’s exhibited more sensitive judging to item ordering, with a rAcc of .36 compared to .18 (TRs) and .32 (URs), and they also had more variability (81.47, SD = 9.27) in their cutoff scores than the URs (81.91, SD = 7.41). The SMEs had less variability in their cutoff scores than the TRs (81.77, SD = 10.97) but were significantly more accurate in their judging indices. These factors, results from the post exercise polling of the judges that no single heuristic was followed, and the variability in the item response curve suggest that hovering systematically around the 80% value did not occur.

Rater analysis was evaluated to examine whether eliminating judges with low rater-total correlations would affect the accuracy and reliability of the judges’ indices.
The results of this study did not support the Maurer et al. (1991) rater analysis technique.

SMEs in the public sector are usually supervisory level personnel and promotion methods differ in private and public sector jobs. Whereas in private sector jobs, supervisors are more often than not, promoted based on profit related productivity, the public sector uses a Training and Experience (T&E) formula as the promotion criterion. The quantitative nature of T&E scales emphasizes quantity of relevant job duties, more so than quality of job duties when used to promote individuals. Coupled with the entry level nature of the evaluated position, these factors might help explain the small observed difference between the UR and SME groups. Criticism of the SME group is not justified in this study however. All three groups were surprising accurate when their ratings were compared to the true cutoff scores. Additionally, the small $n$ of the SME group possibly contributed to the observed statistical anomalies.

Supervisor cognitive processes during performance ratings of employees may contribute to negative subjective evaluations that conflict with the quantitatively derived essential job functions. These factors may have been incorporated into the prototype leading to a broader base of information for the TR group but also providing confusing information to delineate during the judging process.
This greater heterogeneity found in the TR group could possibly be explained by cognitive processes that occur during the performance rating of minimally competent employees. Performance evaluations of minimally competent individuals are typically negative experiences for both the supervisor and employee. During development of the Minimally Competent Clerk I prototype, some subjective personal feelings may have been injected into the profile. This profile, when presented to the TR group, possibly conflicted with the quantitatively produced KSA’s required of a minimally competent worker, this lead to some confusion on their part, and resulted in greater variance and lower accuracy of their ratings.

The greater homogeneity exhibited by the SME’s, though their subjective feelings were undoubtedly present, was probably due to their practical experience mitigating these same subjective feelings. On the other hand, the UR group was presented only the concrete Clerk I KSA’s based on a job analysis, and their homogeneity might be attributed to the objective nature of the KSA’s and essential job functions they received. This suggests that removing subjective evaluations from a prototype profile given to raters may enhance their accuracy.

Limitations

Some limitations of the present study warrant discussion. First, the small number of SMEs is a common
problem associated with research efforts in an applied setting. Justifying removal of highly trained personnel to participate in research, is met in many cases, with resistance from upper management. Second, the small number of SMEs also came from diverse departmental backgrounds. Although the essential job functions were similar for each department, the time invested in, and time needed to complete job duties varied for a particular clerical job in each department.

Finally, the word prototype implies a model, usually a positive or perfect one. A minimally competent employee, usually a negative description, suggests someone who could very easily be incompetent. Combining these two polarized concepts, as presented to the TR group, may have lead to confusion, and may have contributed to some of the inconsistent findings for this group. Substituting the word "profile" might have more neutral effects on the TR raters.

Future Research

In addition to removing subjective evaluations from the prototype, additional research might focus on the cognitive processes occurring during employee performance evaluations and also the differing cognitive judgement strategies in use by the raters. Although the consistent cutoff scores might be explained by the entry level, non-technical skill requirements of the position, there appear to be some personal feelings integrated in the prototype. These
personal feelings should be discouraged during job evaluations. Additionally, the performance by the control group further reinforces how imperative it is to conduct a thorough job analysis, since that may be all that is necessary to establish accurate cutting scores in lower level non-technical professions.

Other areas of future research to consider might focus on replication of this study in private sector jobs and also high-technical positions. The dynamics revolving around these two environments may produce starkly different results.
APPENDIX A

SECURITY AGREEMENT LETTER
INTER-OFFICE CORRESPONDENCE

TO Charles Baker
FROM Dick Hodapp
Date December 21, 1992
SUBJECT Security for Exam Used in Thesis Project

I am pleased to be able to assist you in conducting the research for your thesis project. There are, however, several conditions that you must understand and agree to before I will be able to let you take possession of the City of Fort Worth's General Clerical Abilities Exam. This test is extremely valuable to the City of Fort Worth. A considerable expenditure in City money and manpower was expended in its development and validation. The maintenance of security for this test is our primary concern in letting you use it in your thesis research. To insure the security of this test, I ask that you agree to the following:

- No reproduction of the test is allowed. We will provide you with the copies you need for your research. In your thesis you may provide the summary sheet describing the test but you may not print the actual test. It is typical to include all research material in an appendix. Test manuals normally contain a statement to the effect that "The test has not been included in this report for the purpose of test security." You will have to use the same format in your thesis.

- You must maintain possession of the test in your physical presence at all times. You may allow your thesis committee and the university's research committee to review the test but you must be physically present at all times. No copies of the test may be made. We will provide you with a summary sheet that describes the test and each of its subparts. When the test is used in the research, you must be physically present and personally account for each test used.

While these requirements may seem a little heavy-handed, please keep in mind that this is a working test which has great utility for the City of Fort Worth. If its security were lost, we would have to discontinue its use.

Please provide me with your written agreement to these conditions and we will assist you in any way we can with your research.
APPENDIX B

SECURITY AGREEMENT RESPONSE
January 29, 1993

Dick Hodapp  
Personnel Administrator  
City of Fort Worth, Texas

Re: Agreement to Test Security Measures Used for Thesis Project

Dear Dick,

I appreciate all the help you have provided me toward completion of my thesis project. In response to your memo of December 21, 1992, I agree to all security measures requested therein. I look forward to moving forward with this project, and again thanks for your help.

Sincerely,

Charles E. Baker

CC: Doug Johnson, Ph.D.
APPENDIX C

SUBJECT INFORMATION FORM
SUBJECT INFORMATION

Today's Date: 

Sex:  M  F  Age: ______  Race: ______

Years of Education/Degree: ______

Major: ________________

Current Occupation: ________________

Years of Work Experience: ______

Years of Clerical/Secretarial Experience: ______

Years of Clerical/Secretary Supervisory Experience: ______
APPENDIX D

INFORMED CONSENT (UNIVERSITY SUBJECTS)
USE OF HUMAN SUBJECTS
INFORMED CONSENT

I, _______________________, agree to participate in a study of personnel selection methods conducted by CHARLES BAKER at the University of North Texas. The purpose of this study is educational and is a training requirement of the above named student for graduate level independent research at The University of North Texas. As such, I understand that the results of this exercise WILL NOT be explained to me. I further understand that the results of this evaluation are confidential and will be used ONLY for educational purposes.

As a participant in this study, I understand that I will be expected to participate in a number of experimental tasks which may include the completion of forms, checklists, simulated exercises, and questionnaires relating to my knowledge, attitudes, and behavior, and the occasional observation of my activities. These instruments may include behavioral logs or diaries, attitudinal surveys, activity checklists, and informational quizzes. In addition, I have been told that this is a project that requires no further participation past the date the exercise is conducted. If the need arises, and I am asked to continue participation, I will be told exactly what further participation will entail.

I have been informed that any information obtained in this study will be recorded with a code number that will allow CHARLES BAKER to determine my identity. At the conclusion of this study the key that relates my name with my assigned code number will be destroyed. Under this condition, I agree that any information obtained from this research may be used in any way thought best for publication or education.

I understand that there is no personal risk or discomfort directly involved with this research and that I am free to withdraw my consent and discontinue participation in this study at any time. A decision to withdraw from the study will not affect my standing in any class I am currently enrolled. If I have any questions or problems that arise in connection with my participation in this study, I should contact CHARLES BAKER, the project director at 817/232-0019 (home).

_________________________________________  ______________________________________
(Participant)  (Date)

_________________________________________  ______________________________________
(Investigator)  (Date)

THIS PROJECT HAS BEEN REVIEWED BY THE UNIVERSITY OF NORTH TEXAS COMMITTEE FOR THE PROTECTION OF HUMAN SUBJECTS (Phone: 565-3940).
APPENDIX E

INFORMED CONSENT (CITY OF FORT WORTH, TX)
INFORMED CONSENT

NAME OF PARTICIPANT

I hereby give consent to Charles Baker to perform a personnel selection method study.

I understand that participation is voluntary and that confidentiality will be maintained with regard to individual participant names. All other scoring information will be made available to the Testing Division of the Fort Worth Personnel Department.

I understand that there is no personal risk or discomfort directly involved with this research, the procedure or treatment to be performed is investigational, and that I may withdraw my consent at any time without prejudice or penalty.

________________________
DATE

________________________
PARTICIPANT SIGNATURE
APPENDIX F

DESCRIPTION OF CLERICAL EXAM
The General Clerical Abilities Exam is administered by the City of Fort Worth Personnel Department, Testing Division. The exam consists of nine sections with a total of 125 multiple choice questions. The time allowed to complete the exam is two hours. You may not use calculators, dictionaries, or any materials other than those provided for you by the examiner.

Following is a brief description of each section of the exam:

1. **Spelling** (10 questions) - The spelling questions are designed to measure your ability to recognize words that are spelled correctly. Given four different spellings of the same word, you are to select the choice with the correct spelling.

2. **Vocabulary** (15 questions) - The vocabulary questions are designed to measure your ability to understand the meaning of words. Given four different words, you are to select the word that means the same as the underlined word.

3. **Office Practices** (25 questions) - The office practices questions deal with a variety of situations which require a knowledge of general office practices. After reading each question, you are to choose the answer which reflects the best way to handle the situation.

4. **Filing** (10 questions) - The filing questions require a knowledge of alphabetical and numerical filing. Given a list of four names or numbers, you must determine which item would be filed first, second, third, or fourth, depending on the directions for each question.

5. **Forms Checking/Coding** (15 questions) - The forms checking questions (Part A - 10 questions) examine your ability to check a form for completeness and accuracy, after you are given instructions and information on how the form should be completed. The coding questions (Part B - 5 questions) are designed to measure your ability to code information accurately from tables provided on the exam.

6. **Reading Comprehension** (10 questions) - The reading comprehension questions are designed to measure your ability to understand what you read. In Part A (5 questions), you are to read passage and find the one underlined word that does not fit the meaning of the paragraph. Part B (5 questions) requires you to read a memo and answer questions about that memo.

7. **Checking** (20 questions) - The checking questions examine your ability to check names and numbers for errors. You are given two sets of information with three lines in each set. You are to determine if both sets are alike or if there are one, two, or three errors.

8. **Language Usage** (10 questions) - The language usage questions are designed to measure your knowledge of proper grammar and English usage. Given four sentences, you are to select the one sentence which is the most correct from the standpoint of grammar and English usage.

9. **Math Reasoning** (10 questions) - The math reasoning questions are designed to measure your ability to solve math reasoning word problems involving operations such as addition, subtraction, multiplication, division and calculating percentages. After reading and working each word problem, you are to select the correct answer from the answers provided.
APPENDIX G

ANGOFF PERCENTAGE SHEETS
Percentage Sheet

"Minimally competent" people are those people who have just enough of the Final Knowledge, Skills, and Abilities (KSA's) to be able to do the job of Clerical Worker I. If they had less of these knowledge and abilities, they could not succeed as a Clerical Worker I.

**KNOWLEDGE, SKILLS, AND ABILITIES:** Knowledge of English grammar and usage, vocabulary, spelling, punctuation, mathematical computation, using the four basic arithmetic operations and correct telephone etiquette. Knowledge of postage machine, present operations, third class postal requisitions and post office procedures. Skilled in operating computer, postage equipment inserter, folder, and scales. Ability to understand and follow written and/or oral instructions; perform routine and repetitive work; perform mathematical operations; read and copy information without errors; establish and maintain an effective working relationship with other employees and the general public.

For each question, circle the percentages of the "minimally competent" people you think would get the answer right.

1. 10%  20%  30%  40%  50%  60%  70%  80%  90%
2. 10%  20%  30%  40%  50%  60%  70%  80%  90%
3. 10%  20%  30%  40%  50%  60%  70%  80%  90%
4. 10%  20%  30%  40%  50%  60%  70%  80%  90%
5. 10%  20%  30%  40%  50%  60%  70%  80%  90%
6. 10%  20%  30%  40%  50%  60%  70%  80%  90%
7. 10%  20%  30%  40%  50%  60%  70%  80%  90%
8. 10%  20%  30%  40%  50%  60%  70%  80%  90%
9. 10%  20%  30%  40%  50%  60%  70%  80%  90%
10. 10%  20%  30%  40%  50%  60%  70%  80%  90%
11. 10%  20%  30%  40%  50%  60%  70%  80%  90%
12. 10%  20%  30%  40%  50%  60%  70%  80%  90%
13. 10%  20%  30%  40%  50%  60%  70%  80%  90%
14. 10%  20%  30%  40%  50%  60%  70%  80%  90%
15. 10%  20%  30%  40%  50%  60%  70%  80%  90%
16. 10%  20%  30%  40%  50%  60%  70%  80%  90%
17. 10%  20%  30%  40%  50%  60%  70%  80%  90%
18. 10%  20%  30%  40%  50%  60%  70%  80%  90%
19. 10%  20%  30%  40%  50%  60%  70%  80%  90%
<table>
<thead>
<tr>
<th></th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>60%</td>
<td>70%</td>
<td>80%</td>
<td>90%</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>60%</td>
<td>70%</td>
<td>80%</td>
<td>90%</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>104.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>106.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>107.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>108.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>111.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>112.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>113.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>114.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>116.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>117.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>118.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>119.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>121.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>122.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>123.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>124.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>125.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
DIRECTIONS FOR COMPLETING PERCENTAGE SHEET

Based on your previously developed Clerk I "prototype" please:

1. Follow the guidelines for "minimally competent clerical employee" as defined on the percentage sheets.

2. Mentally develop a "minimally competent clerical prototype" based on your expert opinion and relevant CRITICAL INCIDENTS.

3. Mentally determine what makes them marginal employees based on specific incidents as you evaluate each item.
APPENDIX I

SME MINIMALLY COMPETENT CLERK I EXERCISE
PROTOTYPE OF A MINIMALLY COMPETENT CLERK I

DIRECTIONS: Please complete the following questions based on your department and observed CRITICAL INCIDENTS.

I. FILING

1. When filing materials such as letters, correspondence, index cards, reports, and contracts, a minimally competent Clerk I is capable of:
   a. Misfiling no more than ____ to ____ items in a six month period.
   b. Accurately filing a minimum of ____ to ____ items per day.
   c. Accurately filing the requested materials within ____ to ____ hours of the request.

2. Generally, when filing materials, a person would not be successful as a Clerk I if:

II. MANAGEMENT OF RECORDS

1. While maintaining, retrieving, and opening records for work utilization, a minimally competent Clerk I is capable of:
   a. Performing these duties with no more than ____ to ____ errors in a six month period.
   b. Accurately performing a minimum of ____ to ____ of these duties per day.
   c. Accurately performing the requested duties within ____ to ____ hours of the request.

2. While completing routine forms, a minimally competent Clerk I is capable of:
   a. Completing these forms with no more than ____ to ____ errors in a six month period.
   b. Accurately completing a minimum of ____ to ____ forms per day.
   c. Accurately completing these forms within ____ to ____ hours of the request.
3. Generally, when managing records or completing routine forms, a person would not be successful as a Clerk I if:

III. MAIL HANDLING

1. When receiving, picking up, processing, sorting, opening and distributing mail, or preparing mail for post mark, a minimally competent Clerk I is capable of:

   a. Performing these duties with no more than _____ to _____ errors in a six month period.
   b. Accurately performing a minimum of _____ to _____ of these duties per day.
   c. Accurately performing the requested duties within _____ to _____ hours of the request.

2. Generally, when handling mail, a person would not be successful as a Clerk I if:

IV. INFORMATION TO THE GENERAL PUBLIC

1. When providing information to the public in person or by phone, a minimally competent Clerk I is capable of:

   a. Providing information with no more than _____ to _____ errors or complaints from the public in a six month period.
   b. Accurately providing _____ to _____ bits of information to the public per day.
   c. Accurately providing the requested information to the public within _____ to _____ hours of the request.

2. Generally, when providing information to the public, a person would not be successful as a Clerk I if:
V. REPORT HANDLING

1. When assembling reports, (making copies, collating, and stapling) a minimally competent Clerk I is capable of:
   
   a. Assembling reports with no more than ____ to ____ errors in a six month period of time.
   b. Accurately assembling ____ to ____ reports per day.
   c. Accurately providing the requested assembled report within ____ to ____ hours of the request.

2. Generally, when assembling reports, a person would not be successful as a Clerk I if:

VI. MONEY EXCHANGE

1. When receiving money and issuing receipts, a minimally competent Clerk I is capable of:
   
   a. Exchanging money and issuing receipts with no more than ____ to ____ mistakes in a six month period.
   b. Accurately conducting ____ to ____ monetary transactions and issuing receipts per day.
   c. Accurately handling the money transactions and issuing receipts within ____ to ____ hours of the request.

2. Generally, when conducting money transactions, a person would not be successful as a Clerk I if:
APPENDIX J

SME COMPOSITE PROTOTYPE
### MINIMALLY COMPETENT CLERK I

#### MISFILING

<table>
<thead>
<tr>
<th>Quality</th>
<th>Quantity</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>300/day</td>
<td>7 hrs</td>
</tr>
<tr>
<td>1%</td>
<td>600</td>
<td>10</td>
</tr>
<tr>
<td>1%</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>1%</td>
<td>230</td>
<td>1</td>
</tr>
<tr>
<td>10%</td>
<td>1500</td>
<td>7</td>
</tr>
<tr>
<td>@*3%</td>
<td>531 ave.</td>
<td>5 ave.</td>
</tr>
</tbody>
</table>

Unable to follow instructions.

***Dyslexia***

Poor organizational skills; not adjusting numbers or letters.

If they only have limited rather than close supervision.

Have trouble distinguishing between various types of documents.

Low comfort level with the alphabet or numbers.

Poor ability to communicate with coworkers and supervisor.

Poor reasoning skills.

Inability to follow written/oral instructions.

### MANAGEMENT OF RECORDS

<table>
<thead>
<tr>
<th>Quality</th>
<th>Quantity</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>15</td>
<td>2 hrs</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>55</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>100</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>37 ave</td>
<td>16 ave</td>
<td>3 ave</td>
</tr>
</tbody>
</table>

Completing Forms

<table>
<thead>
<tr>
<th>Quality</th>
<th>Quantity</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5</td>
<td>2 hrs</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>20</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>200</td>
<td>6</td>
</tr>
<tr>
<td>50</td>
<td>200</td>
<td>6</td>
</tr>
<tr>
<td>17 ave.</td>
<td>87 ave.</td>
<td>5 ave.</td>
</tr>
</tbody>
</table>

Lack knowledge of Personal Computer environment.

Inability to read a set of instructions.
Inability to retain oral instructions.
Lack basic math skills.
Unable to read and transfer information into documents.
Inability to log what they managed and completed versus just working with files, i.e., a production report.

### MAIL HANDLING

<table>
<thead>
<tr>
<th>Quality</th>
<th>Quantity</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>75</td>
<td>5</td>
</tr>
<tr>
<td>50</td>
<td>400</td>
<td>6</td>
</tr>
<tr>
<td>19 ave.</td>
<td>101 ave.</td>
<td>3 ave.</td>
</tr>
</tbody>
</table>

Inability to follow oral/written directions.
Poor organization skills (cannot distinguish between interoffice and outer-office mail) as it relates to the proper mail destination.
Inability to ask proper procedural questions and then follow the given answers.
Inability to merge documents (for form letter purposes).
Inability to match variables (person who receives a ticket versus a victim of a crime)

***Manual dexterity***

### INFORMATION TO PUBLIC

<table>
<thead>
<tr>
<th>Quality</th>
<th>Quantity</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10</td>
<td>5 minutes</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>75</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>500</td>
<td>10</td>
</tr>
<tr>
<td>4 ave.</td>
<td>129 ave.</td>
<td>12 minute ave.</td>
</tr>
</tbody>
</table>

Not courteous to public.
Lack good grammar skills and grammar usage during phone or personal public contact.
Have poor communication skills when interacting with an irate or uninformed customer.
Lack knowledge of telephone procedures; lack knowledge of department information available to public.
Have poor phone etiquette.
Has little patience with the public.
REPORT HANDLING

<table>
<thead>
<tr>
<th>Quality</th>
<th>Quantity</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2</td>
<td>1 hour</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>500</td>
<td>6</td>
</tr>
<tr>
<td>8 ave.</td>
<td>108 ave.</td>
<td>2 hrs ave.</td>
</tr>
</tbody>
</table>

Inability to follow oral/written instructions.
Poor organizational skills.
Lack ability to operate copier/office equipment efficiently.

MONEY EXCHANGE

<table>
<thead>
<tr>
<th>Quality</th>
<th>Quantity</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>15</td>
<td>10 minutes</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>240</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3 ave.</td>
<td>53 ave.</td>
<td>19 minute ave.</td>
</tr>
</tbody>
</table>

Poor math skills (returning incorrect change).
Inability to do general bookkeeping functions.
Inability to follow established procedures.

GENERAL COMMENTS

Ask Testing Administrator to:

- Raise cutoff scores.
- Grade by subtest area.
- Develop a ten key test.
- Develop a CRT and PC test.
- Develop a software test (Wordperfect 5.1)
- Have new typewriters for test taking.
- Have five (5) minute breaks during tests (in the room).
APPENDIX K

PROTOTYPE GIVEN TO TRAINED RATERS
CLERK I PROFILE

DIRECTIONS: Please read the following information carefully. If there are any questions please ask the investigator.

FILING

When filing materials such as letters, correspondence, index cards, reports, and contracts, a minimally competent Clerk I is capable of: Accurately filing a minimum of 531 items/day within 5 hours of being requested to perform this task. They would misfile no more than 50 items in a 6 month period.

Generally, when filing materials, a person would not be successful as a Clerk I if they: Have trouble distinguishing between various types of documents, have a low comfort level with the alphabet or numbers, and are unable to follow written/oral instructions.

MANAGEMENT OF RECORDS

While maintaining, retrieving, and opening records for work utilization, a minimally competent Clerk I is capable of: Accurately performing a minimum of 16 of these duties/day within 3 hours of being requested to perform this task. They would have no more than 37 errors in a 6 month period.

While completing routine forms, a minimally competent Clerk I is capable of: Accurately completing a minimum of 87 forms/day within 5 hours of being requested to perform this task. They would have no more than 17 errors in a six month period.

Generally, when managing records or completing routine forms, a person would not be successful as a Clerk I if they: Lack knowledge of the Personal Computer environment, lack basic math skills, and are unable to read and transfer information into documents.

MAIL HANDLING

When receiving, picking up, processing, sorting, opening and distributing mail, or preparing mail for post mark, a minimally competent Clerk I is capable of: Accurately performing a minimum of 101 of these duties/day within 3 hours of being requested to perform this task. They would have no more than 19 errors in a six month period.

Generally, when handling mail, a person would not be successful as a Clerk I if they: Have poor organization skills (cannot distinguish between interoffice and outer-office mail) as it relates to the proper mail destination, are unable to merge documents (for form letter purposes), and are unable to match variables (person who receives a ticket versus a victim of a crime).
INFORMATION TO GENERAL PUBLIC

When providing information to the public in person or by phone, a minimally competent Clerk I is capable of: Accurately providing 129 bits of information to the public/day within 12 minutes of the request from the public. They would have no more than 4 errors or complaints from the public in a six month period.

Generally, when providing information to the public, a person would not be successful as a Clerk I if they: Are not courteous to the public, lack good grammar skills and grammar usage during phone or personal public contact, have poor communication skills when interacting with an irate or uninformed customer, lack knowledge of telephone procedures, lack knowledge of department information available to public, have poor phone etiquette, and have little patience with the public.

REPORT HANDLING

When assembling reports, (making copies, collating, and stapling) a minimally competent Clerk I is capable of: Accurately assembling 108 reports/day within 2 hours of being requested to perform this duty. They would have no more than 8 errors in a six month period.

Generally, when assembling reports, a person would not be successful as a Clerk I if they: Are unable to follow oral/written instructions, have poor organizational skills, and lack the ability to operate copier/office equipment efficiently.

MONEY EXCHANGE

When receiving money and issuing receipts, a minimally competent Clerk I is capable of: Accurately conducting 53 monetary transactions and issuing receipts/day. They should complete the transaction within 19 minutes and have no more than 3 mistakes in a 6 month period.

Generally, when conducting money transactions, a person would not be successful as a Clerk I if they: Have poor math skills (returning incorrect change), are unable to do general bookkeeping functions, and are unable to follow established procedures.
APPENDIX L

TRAINED RATER INSTRUCTION SHEET
DIRECTIONS FOR COMPLETING PERCENTAGE SHEET

Based on the previously provided Clerk I "prototype" information please:

1. Follow the guidelines for "minimally competent clerical employee" as defined on the percentage sheets.

2. Mentally develop a "minimally competent clerical prototype" based on your opinion and the previously provided Clerk I information.

3. Mentally determine what makes them marginal employees based on the previously provided information for Clerk I as you evaluate each item.
APPENDIX M

RATER ITEM RESPONSES
GROUP 1 (SMEs)

I1  Test Item 1

Valid cases 5  Missing cases 0

I2  Test Item 2

Valid cases 5  Missing cases 0

I3  Test Item 3

Valid cases 5  Missing cases 0

I4  Test Item 4

Valid cases 5  Missing cases 0

I5  Test Item 5

Valid cases 5  Missing cases 0
16 Test item 6

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Valid cases 5 Missing cases 0

17 Test item 7

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Valid cases 5 Missing cases 0

18 Test item 8

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Valid cases 5 Missing cases 0

19 Test item 9

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Valid cases 5 Missing cases 0

20 Test item 10

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Valid cases 5 Missing cases 0
Test item 16

Valid cases: 5  
Missing cases: 0

Test item 17

Valid cases: 5  
Missing cases: 0

Test item 18

Valid cases: 5  
Missing cases: 0

Test item 19

Valid cases: 5  
Missing cases: 0

Test item 20

Valid cases: 5  
Missing cases: 0
121 Test item 21

122 Test item 22

123 Test item 23

124 Test item 24

125 Test item 25
136 test item 36

Valid cases 5   Missing cases 0

137 test item 37

Valid cases 5   Missing cases 0

138 test item 38

Valid cases 5   Missing cases 0

139 test item 39

Valid cases 5   Missing cases 0

140 test item 40

Valid cases 5   Missing cases 0
141 Test item 41

Valid cases: 5  Missing cases: 0

142 Test item 42

Valid cases: 5  Missing cases: 0

143 Test item 43

Valid cases: 5  Missing cases: 0

144 Test item 44

Valid cases: 5  Missing cases: 0

145 Test item 45

Valid cases: 5  Missing cases: 0
151  test item 51

Valid cases 5  Missing cases 0

152  test item 52

Valid cases 5  Missing cases 0

153  test item 53

Valid cases 5  Missing cases 0

154  test item 54

Valid cases 5  Missing cases 0

155  test item 55

Valid cases 5  Missing cases 0
161  test item 61

Valid cases  5  Missing cases  0

162  test item 62

Valid cases  5  Missing cases  0

163  test item 63

Valid cases  5  Missing cases  0

164  test item 64

Valid cases  5  Missing cases  0

165  test item 65

Valid cases  5  Missing cases  0
166  test item 66

Valid cases 5  Missing cases 0

167  test item 67

Valid cases 5  Missing cases 0

168  test item 68

Valid cases 5  Missing cases 0

169  test item 69

Valid cases 5  Missing cases 0

170  test item 70

Valid cases 5  Missing cases 0
171  test item 71

Valid cases 5  Missing cases 0

172  test item 72

Valid cases 5  Missing cases 0

173  test item 73

Valid cases 5  Missing cases 0

174  test item 74

Valid cases 5  Missing cases 0

175  test item 75

Valid cases 5  Missing cases 0
176 test item 76

Valid cases: 5  Missing cases: 0

177 test item 77

Valid cases: 5  Missing cases: 0

178 test item 78

Valid cases: 5  Missing cases: 0

179 test item 79

Valid cases: 5  Missing cases: 0

180 test item 80

Valid cases: 5  Missing cases: 0
I81  Test item 81

<table>
<thead>
<tr>
<th>Valid cases</th>
<th>Missing cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

I82  Test item 82

<table>
<thead>
<tr>
<th>Valid cases</th>
<th>Missing cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

I83  Test item 83

<table>
<thead>
<tr>
<th>Valid cases</th>
<th>Missing cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

I84  Test item 84

<table>
<thead>
<tr>
<th>Valid cases</th>
<th>Missing cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

I85  Test item 85

<table>
<thead>
<tr>
<th>Valid cases</th>
<th>Missing cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>
186 test item 86

Valid cases 5  Missing cases 0

187 test item 87

Valid cases 5  Missing cases 0

188 test item 88

Valid cases 5  Missing cases 0

189 test item 89

Valid cases 5  Missing cases 0

190 test item 90

Valid cases 5  Missing cases 0
191  test item 91

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 5  Missing cases 0

192  test item 92

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 5  Missing cases 0

193  test item 93

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 5  Missing cases 0

194  test item 94

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 5  Missing cases 0

195  TRest item 95

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 5  Missing cases 0
test item 96

Valid cases 5  Missing cases 0

Valid cases 5  Missing cases 0

Valid cases 5  Missing cases 0

Valid cases 5  Missing cases 0

Valid cases 5  Missing cases 0
<table>
<thead>
<tr>
<th>Test Item</th>
<th>Valid Cases</th>
<th>Missing Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>102</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>103</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>104</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>105</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>
1116  test item 116

Valid cases 5  Missing cases 0

1117  test item 117

Valid cases 5  Missing cases 0

1118  test item 118

Valid cases 5  Missing cases 0

1119  test item 119

Valid cases 5  Missing cases 0

1120  Test item 120

Valid cases 5  Missing cases 0
I121  test item 121

Valid cases  5  Missing cases  0

I122  test item 122

Valid cases  5  Missing cases  0

I123  test item 123

Valid cases  5  Missing cases  0

I124  test item 124

Valid cases  5  Missing cases  0

I125  test item 125

Valid cases  5  Missing cases  0
GROUP 2 (TRs)

11 Test Item 1

Valid cases 20  Missing cases 0

12 Test Item 2

Valid cases 20  Missing cases 0

13 Test Item 3

Valid cases 20  Missing cases 0

14 Test Item 4

Valid cases 20  Missing cases 0
I5  Test item 5

Valid cases 20  Missing cases 0

I6  Test item 6

Valid cases 20  Missing cases 0

I7  Test item 7

Valid cases 20  Missing cases 0

I8  Test item 8

Valid cases 20  Missing cases 0
121 Test item 21

Valid cases 20  Missing cases 0

122 Test item 22

Valid cases 20  Missing cases 0

123 Test item 23

Valid cases 20  Missing cases 0

124 Test item 24

Valid cases 20  Missing cases 0
Valid cases 20  Missing cases 0
I33 test item 33

Valid cases 20  Missing cases 0

I34 test item 34

Valid cases 20  Missing cases 0

I35 test item 35

Valid cases 20  Missing cases 0

I36 test item 36

Valid cases 20  Missing cases 0
141 Test item 41

Valid cases 20 Missing cases 0

142 test item 42

Valid cases 20 Missing cases 0

143 test item 43

Valid cases 20 Missing cases 0

144 test item 44

Valid cases 20 Missing cases 0
153  test item 53

Valid cases  20  Missing cases  0

154  test item 54

Valid cases  20  Missing cases  0

155  test item 55

Valid cases  20  Missing cases  0

156  test item 56

Valid cases  20  Missing cases  0
Test item 81

```
1  1  4  4  4  4  5
0  1  2  3  4  5
```

Valid cases 20  Missing cases 0

Test item 82

```
1  3  2  2  6  6
0  2  4  6  8  10
```

Valid cases 20  Missing cases 0

Test item 83

```
1  4  3  5  7
0  2  4  6  8  10
```

Valid cases 20  Missing cases 0

Test item 84

```
1  4  2  7  7
0  2  4  6  8  10
```

Valid cases 20  Missing cases 0
Valid cases 20  Missing cases  0

Valid cases 20  Missing cases  0

Valid cases 20  Missing cases  0

Valid cases 20  Missing cases  0

Valid cases 20  Missing cases  0
Valid cases: 20  Missing cases: 0
GROUP 3 (UTRs)

11 Test Item 1

Valid cases 20 Missing cases 0

12 Test Item 2

Valid cases 20 Missing cases 0

13 Test Item 3

Valid cases 20 Missing cases 0

14 Test Item 4

Valid cases 20 Missing cases 0
15  Test item 5

Valid cases 20  Missing cases 0

16  Test item 6

Valid cases 20  Missing cases 0

17  Test item 7

Valid cases 20  Missing cases 0

18  Test item 8

Valid cases 20  Missing cases 0
Test item 13

Valid cases 20  Missing cases 0

Test item 14

Valid cases 20  Missing cases 0

Test item 15

Valid cases 20  Missing cases 0

Test item 16

Valid cases 20  Missing cases 0
129  Test item 29

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 20  Missing cases 0

130  Test item 30

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 20  Missing cases 0

131  Test item 31

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 20  Missing cases 0

132  Test item 32

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Valid cases 20  Missing cases 0
133  test item 33

Valid cases  20  Missing cases  0

134  test item 34

Valid cases  20  Missing cases  0

135  test item 35

Valid cases  20  Missing cases  0

136  test item 36

Valid cases  20  Missing cases  0
169 test item 69

Valid cases 20 Missing cases 0

170 test item 70

Valid cases 20 Missing cases 0

171 test item 71

Valid cases 20 Missing cases 0

172 test item 72

Valid cases 20 Missing cases 0
I102  test item 102

Valid cases 20  Missing cases 0

I103  test item 103

Valid cases 20  Missing cases 0

I104  test item 104

Valid cases 20  Missing cases 0

I105  test item 105

Valid cases 20  Missing cases 0
### Test Item 122

| Valid cases | 20 | Missing cases | 0 |

<table>
<thead>
<tr>
<th>Score</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>2</td>
</tr>
<tr>
<td>1.4</td>
<td>4</td>
</tr>
<tr>
<td>1.5</td>
<td>4</td>
</tr>
<tr>
<td>1.6</td>
<td>3</td>
</tr>
<tr>
<td>1.7</td>
<td>3</td>
</tr>
</tbody>
</table>

### Test Item 123

| Valid cases | 20 | Missing cases | 0 |

<table>
<thead>
<tr>
<th>Score</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>1</td>
</tr>
<tr>
<td>1.3</td>
<td>3</td>
</tr>
<tr>
<td>1.4</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>6</td>
</tr>
<tr>
<td>1.6</td>
<td>2</td>
</tr>
<tr>
<td>1.7</td>
<td>1</td>
</tr>
</tbody>
</table>

### Test Item 124

| Valid cases | 20 | Missing cases | 0 |

<table>
<thead>
<tr>
<th>Score</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3</td>
<td>2</td>
</tr>
<tr>
<td>1.4</td>
<td>3</td>
</tr>
<tr>
<td>1.5</td>
<td>5</td>
</tr>
<tr>
<td>1.6</td>
<td>3</td>
</tr>
<tr>
<td>1.7</td>
<td>4</td>
</tr>
<tr>
<td>1.8</td>
<td>3</td>
</tr>
</tbody>
</table>

### Test Item 125

| Valid cases | 20 | Missing cases | 0 |

<table>
<thead>
<tr>
<th>Score</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>1</td>
</tr>
<tr>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>1.4</td>
<td>2</td>
</tr>
<tr>
<td>1.5</td>
<td>3</td>
</tr>
<tr>
<td>1.6</td>
<td>2</td>
</tr>
<tr>
<td>1.7</td>
<td>1</td>
</tr>
<tr>
<td>1.8</td>
<td>4</td>
</tr>
</tbody>
</table>

Valid cases | 20 | Missing cases | 0
REFERENCES


